

BERKSON'S METHOD VS. MAXIMUM LIKELIHOOD
IN THE LOGISTIC REGRESSION MODEL

by

RHONDA JENINE UMALI RAGUAL
B.S., UNIVERSITY OF THE PHILIPPINES AT LOS BANOS, 1983

A MASTER'S REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1988

Approved by:

S. K. Perry
Major Professor

LD
2668
R4
STW
1988
R33
C.2

Al1207 312077

TABLE OF CONTENTS	Page
I. Introduction.....	1
II. Derivations	
A. Maximum Likelihood Estimator.....	3
B. Berkson's Estimator.....	5
III. Review of Literature.....	7
IV. Methodology.....	9
V. Results and Discussion.....	11
A. Maximum Likelihood vs. Berkson's.....	11
1. Biases of $\hat{\beta}_0$ and $\hat{\beta}_1$	11
2. Mean Square Errors of $\hat{\beta}_0$ and $\hat{\beta}_1$	11
3. Variances of $\hat{\beta}_0$ and $\hat{\beta}_1$	12
4. Mean Square Errors of $\hat{p}$	12
B. Berkson Estimator in 9 Designs.....	13
C. Maximum Likelihood Estimator in 9 design.....	13
D. Comparison of Designs with Equal Sample Sizes.....	14
VI. Conclusions and Recommendations.....	72
VII. Bibliography.....	73
VIII. Appendices.....	74
A. Gauss Program: Generation of Data.....	74
B. Gauss Program: Berkson's Estimator.....	75
C. SAS Program: Maximum Likelihood Estimator.....	76

I. INTRODUCTION

Several experimental situations involve administering a stimulus to an experimental unit and then observing a binary response. Consider the following case : An experimenter can apply a stimulus at various levels. Subjects are assigned at random to different levels of the stimulus and for each subject, a binary response is recorded. From this set of observations, one determines a model to describe the relationship between the stimulus and the proportion responding to the stimulus. The model is called a dose-response model. These types of models can be applied to a lot of practical problems such as those presented by Chatterjee(1977), Finney(1971), and Milliken(1982). They are notably common in the field of pharmacology, in bioassay, where the levels of the stimulus may represent different doses of a drug or pesticide and the binary response is either death or survival.

Suppose a pesticide is tested at k different levels. At the level of dosage X_i , let m_i be the number of insects that die out of a total n_i exposed to the pesticide ($i = 1, 2, \dots, k$). We want to estimate the relationship between dose, X_i , and the proportion dying, P_i .

The stimulus-response relationships have generally been found to be nonlinear. One nonlinear function which has been used to approximate the relationship between the dose, X_i and the response probability, P_i is

$$P_i = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}} \quad (1.1)$$

where β_0 and β_1 are unknown parameters.

This relationship is referred to as the logistic response function and has the shape shown in Fig. 1.1. The corresponding analysis is referred to in the literature as logit analysis. The logistic function is bounded between 0 and 1, and is monotonic.

Two estimators of β_0 and β_1 in the model have been proposed in the literature. One estimator is the maximum likelihood estimator (MLE) and the other is the method proposed by Berkson in 1953. This study aims to further investigate and compare the behaviors of the two estimators with respect to some dose-response designs by computer simulation.

II. DERIVATIONS

As earlier noted, the logistic response function

$$P_i = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}},$$

is commonly used to approximate the relationship between the dose X_i and the response probability P_i . The main concern is to estimate the parameters β_0 and β_1 . Two estimators of the logistic regression coefficients have been proposed in the literature and are derived in this section.

A. Maximum Likelihood Estimator

Let $X_i = i^{\text{th}}$ dose $i = 1, 2, \dots, k$ levels,

n_i = number of experimental units exposed to the stimulus,

m_i = number of experimental units reacting to the stimulus,

$P_i = \frac{m_i}{n_i}$ = observed 'mortality rate'

P_i = 'true mortality rate' and

$Q_i = 1 - P_i$.

It is usually reasonable to assume that m_i is a Binomial(n_i, P_i) random variable. Thus, the probability density function of m_i is

$$f(m_i; P_i, n_i) = \binom{n_i}{m_i} P_i^{m_i} Q_i^{n_i - m_i} \quad (2.1)$$

for $i = 1, \dots, k$.

The log likelihood function of the model is

$$\ln L = \sum_{i=1}^k \ln \binom{n_i}{m_i} + \sum_{i=1}^k m_i \ln P_i + \sum_{i=1}^k (n_i - m_i) \ln Q_i \quad (2.2)$$

$$\text{where } P_i = \frac{e^{\beta_0 + \beta_1 X_i}}{1 + e^{\beta_0 + \beta_1 X_i}}, \quad Q_i = 1 - P_i.$$

Differentiating $\ln L$ with respect to β_0 and β_1 and setting the corresponding derivatives equal to zero, we obtain the normal equations

$$\sum_{i=1}^k n_i (P_i - \hat{P}_i) = 0 \quad (2.3)$$

and

$$\sum_{i=1}^k n_i X_i (P_i - \hat{P}_i) = 0 \quad (2.4)$$

Solving the normal equations by iteration gives the ML estimators for β_0 and β_1 .

B. Berkson's Method

Let X_i , m_i , n_i , p_i , q_i , P_i , Q_i denote the same quantities as in Section II. A. In the following, we outline the derivation of Berkson's estimators of β_0 and β_1 .

It is well-known that

$$p_i \sim AN \left[p_i, \frac{p_i(1-p_i)}{n_i} \right] \quad (2.5)$$

where "-AN" means "asymptotically normally distributed". Then

$$\ln \left[\frac{p_i}{1-p_i} \right] \sim AN \left[\beta_0 + \beta_1 X_i, \frac{1}{n_i p_i (1-p_i)} \right] \quad (2.6)$$

(See for example p.118, Theorem A, Serfling (1980)). Let

$$l_i = \ln \left[\frac{p_i}{1-p_i} \right] \quad \text{for } i=1, \dots, k.$$

Then one can write:

$$l_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (2.7)$$

where $\epsilon_i \sim N(0, (n_i p_i (1-p_i))^{-1})$ for $i=1, 2, \dots, k$,

and $\epsilon_1, \dots, \epsilon_k$ are independent.

One first obtains the weighted least squares estimators of β_0 and β_1 by using (2.7) under the assumption that P_i 's are known. Then replacing P_i by p_i in the weighted least squares estimators of β_0 and

β_1 , one obtains Berkson's estimators of β_0 and β_1 as follows:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^k n_i p_i q_i 1_i X_i^2 - \frac{\sum_{i=1}^k n_i p_i q_i 1_i \sum_{i=1}^k n_i p_i q_i X_i}{\sum_{i=1}^k n_i p_i q_i}}{\sum_{i=1}^k n_i p_i q_i X_i^2 - \frac{(\sum_{i=1}^k n_i p_i q_i X_i)^2}{\sum_{i=1}^k n_i p_i q_i}} \quad (2.8)$$

and

$$\hat{\beta}_0 = \frac{\sum_{i=1}^k n_i p_i q_i 1_i - \hat{\beta}_1 \sum_{i=1}^k n_i p_i q_i X_i}{\sum_{i=1}^k n_i p_i q_i} \quad (2.9)$$

III. REVIEW OF LITERATURE

Other than the logistic model discussed in Chapter 1, an alternative model in which the response function is represented by the cumulative normal distribution function has also been proposed. The cumulative distribution function of the normal distribution has a shape similar to that of the logistic function (see Fig 1.1). The resulting model is called the probit model and is extensively covered by Finney(1971).

Milliken(1982) discussed dose-response analyses in the light of tolerance distributions. Tolerance of the experimental unit defined as the level of the stimulus at which a response first occurs, he described the tolerance distribution as the collection of tolerances for a population of experimental units and that the assumed form of the tolerance distribution determines the type of analysis. Aside from the most commonly used probit and logit analysis which uses the normal and logistic curves as tolerance distributions, respectively, models resulting from the use of exponential and e to the e function as tolerance distributions were also presented.

For the logit analysis, several studies have been undertaken since the maximum likelihood estimator (MLE) and Berkson's method appeared in the literature. The MLE was first advocated for general application by R.A. Fisher. For this particular model though, the solution requires an iterative procedure as opposed to a

computationally simple estimator suggested by Berkson(1953). This is in fact one of the factors which motivated this study of comparison.

Further interest was brought about by some notable results presented in a theoretical study by Davis(1985). One of the interesting implications of her results was that Berkson's method should not be used when the number of design points is large and the average number of observations per design point is small. Furthermore, she pointed out that when the number of design points is large, Berkson's estimator is consistent if the average number of observations per design point is large. These results, which will eventually (in Chapter V) be compared with the outcome of the simulation study, further motivated this area of research.

For the logit transformation, the value of $\ln \left(\frac{p_i}{1-p_i} \right)$ is not defined for $p_i = 0$ or 1. Chatterjee and Price(1977) noted that for fitting a dose-response relationship, such points are usually omitted, as there is considerable degree of uncertainty about the exact dosage at these response levels. Berkson's proposal(1955), which was used in this study, was to use the following rule:

$$p_i = \begin{cases} \frac{1}{2n_i} & , \quad m_i = 0 \\ 1 - \frac{1}{2n_i} & , \quad m_i = n_i \end{cases}$$

where p_i , n_i , and m_i are as defined in Chapter II.

IV. METHODOLOGY

The special case of a simple logit model with one independent variable was examined. For all experiments, the parameters of the logistic function were $\beta_0 = -4.12$ and $\beta_1 = 1.18$. Nine dose-response designs were investigated: three five-dose designs, three ten-dose designs and three fifteen-dose designs. The experiments were conducted with an equal number of observations per cell, $n_i = n$ for all i . In all three main simulations, the cell sizes were $n = 10, 20, 30$. Therefore, the total sample sizes for the five-dose designs were $N = 50, 100, 150$; $N = 100, 200, 300$ for the ten-dose designs and $N = 150, 300, 450$ for the fifteen-dose designs.

For each of the 9 simulations, 500 sets of data were generated through GAUSS(1987) (for the algorithm, see Appendix 1) and were analyzed using the two estimation methods. Here, for the aforementioned β 's ($\beta_0 = -4.12$ and $\beta_1 = 1.18$) and specified X_i 's (k doses

equally spaced on log doses 1-6), $P_i = \frac{\beta_0 + \beta_1 X_i}{1 + e^{\beta_0 + \beta_1 X_i}}$ were computed

and m_i 's were generated from a Binomial(n_i, P_i) random variable where n_i 's were the specified number of observations per dose in the design. The Berkson estimates were obtained using a GAUSS algorithm while that of maximum likelihood were computed using PROC NLIN of SAS(1982).

As a basis for comparison, biases, variances and mean square

errors(MSE) of the estimated regression coefficients ($\hat{\beta}_0, \hat{\beta}_1$) were calculated for each design. In addition, as a measure of how the estimated p 's($\hat{p}$) compare with the actual P 's, the quantity

$$\frac{\sum_{i=1}^k (\hat{p}_i - P_i)^2}{k}$$

was computed for each dose-response data set and averaged over 500 data sets under each design. This in essence, measures on the average, how close the predicted logistic regression curve is from the actual. Throughout the report, the quantity

$$\frac{s}{j=1} \left[\frac{\sum_{i=1}^k (\hat{p}_i - P_i)^2}{k} \right]^2 \quad (4.1)$$

where s = number of data sets(in this case 500)

will be referred to as $MSE(\hat{p})$.

V. RESULTS AND DISCUSSION

A. Maximum Likelihood vs. Berkson's

Simulation estimates of the bias, variance and MSE of the sampling distributions of the estimated $\hat{\beta}_0$ and $\hat{\beta}_1$ are presented in Table 5.1. MSE(p)(see 4.1) are also included in the table.

1. Biases of $\hat{\beta}_0$ and $\hat{\beta}_1$

Simulation results indicate that the Berkson statistic is more biased than the maximum likelihood estimator (MLE). For the small sample size within dose designs ($k = 5, 10, 15; n=10$) the differences were not great. Berkson biases were only .78 to 2.5 times larger than MLE's. These are graphically demonstrated in Fig. 5.1-5.12. Differences were most pronounced in larger sample size designs ($n=20, 30$ under all k), as illustrated in Fig. 5.13-5.32. In these cases, MLE essentially dominated Berkson's (with Berkson bias being as much as 176 times that of MLE) except in the 5 dose, $n=30$ design. Here, MLE bias was about 1.5 times larger than Berkson's. This was the only case where Berkson's fared better than the MLE in terms of bias. Figures 5.33-5.36 suggest though that Berkson's advantage is small.

2. Mean Square Errors of $\hat{\beta}_0$ and $\hat{\beta}_1$

Mean square error results for the regression coefficients (β_0, β_1) were consistent with the performance of the estimators with respect to biases. In this case though, the discrepancies were rather small.

Berkson MSE's were only 7% to 56% higher than MLE's, except again for the 5 dose, n=30 design where MSE of maximum likelihood estimators for β_0 and β_1 were 22% and 38% more than their Berkson counterparts.

3. Variances of $\hat{\beta}_0$ and $\hat{\beta}_1$

The variances, on the other hand, exhibited a slightly different trend. Though not extremely nor uniformly better, the Berkson estimates showed less variability as seen in Table 5.1. This was exhibited in all of the 5 dose designs (n=10, 20, 30) and in the 15 dose, n=20 design where MLE variances were 1 to 30% larger than the corresponding Berkson's. Plots in Fig. 5.29-5.32 particularly reveal this Berkson's rather slight edge over MLE in some designs. However, maximum likelihood still yielded slightly smaller variances for both $\hat{\beta}_0$ and $\hat{\beta}_1$ in the 15 dose, n=30 design (Fig. 5.13-5.16). In all other cases the trend was inconclusive. There was inconsistency in the behavior of the variances within a design. That is, either Berkson's gave lower variance for $\hat{\beta}_0$ and higher variance for $\hat{\beta}_1$ or vice-versa.

4. Mean Square Error of $\hat{p}$

In most cases, experimenters are not particularly interested in the specific values of β_0 and β_1 but are concerned with estimating P_i , which is the actual proportion of experimental units reacting to the stimulus at a certain level of dosage. A quantity which measures the 'combined ability' of $\hat{\beta}_0$ and $\hat{\beta}_1$, in predicting p_i was computed. This

was the $MSE(\hat{p})$ discussed in Chap. IV(4.1). It essentially measures how close, on the average, the predicted logistic regression curve is to the actual curve.

The maximum likelihood estimator showed complete superiority in this respect as evidenced in Fig. 5.37-5.54. Berkson's estimates produced $MSE(\hat{p})$'s which were 3% to 69% higher than the MLE counterparts. It was also observed that the discrepancy increased as the sample size increased within a specified number of doses (k).

B. Berkson Estimator in 9 designs

For a fixed number of doses(k), as the sample size increased the bias decreased, the variances were reduced, and MSE's decreased in all of the designs as anticipated. It is again worth noting that the Berkson estimator was at its best as far as bias was concerned in the 5 dose, 30 observations per dose design.

C. Maximum Likelihood Estimator in 9 designs

As the sample size n increased within a fixed number of doses(k), the improvement in bias, variance, and MSE, that is, in the performance of the estimator in general, was not as great as that in Berkson. This shows the relative stability of the MLE regardless of the design. In fact, for $k=5$ and $k=10$, biases were even smaller when $n=20$ than when $n=30$.

D. Comparison of Designs with Equal Total Sample Sizes N

This part of the analysis is focused on the practical aspect of experimentation. Most often, experimenters are faced with a fixed number of experimental units at hand and the immediate problem is

choosing a proper dose-response design to most closely predict p based on a fitted logistic model. Is one better off using more design points or using more observations per design point?

1. $N=100 : k=5 \ n=20$ vs. $k=10 \ n=10$

The 5 dose, 20 observations per dose design resulted in lower $MSE(\hat{p})$ for both methods as shown in Table 5.1 and Fig. 5.39-5.40, 5.43-5.44. No definite trend was discerned among the $\hat{\beta}_0$'s and $\hat{\beta}_1$'s not only in this case but in the other two comparisons below as well.

2. $N=150 : k=5 \ n=30$ vs. $k=15 \ n=10$

Here, Table 5.1 and Fig. 5.41-5.42, 5.49-5.50 conspicuously reveal the superiority of the 5 dose, 30 observations per dose design over the 15 dose, 10 observations per dose design, in predicting the proportion(p) of experimental units responding to the stimulus.

3. $N=300 : k=10 \ n=30$ vs. $k=15 \ n=20$

The 15-dose, 20 observations per dose design is slightly better in predicting $\hat{p}$ as evidenced by its low $MSE(\hat{p})$ in both designs. Fig. 5.47-5.48, 5.51-5.52 attest to this.

Table 5.1 Simulation Estimates of the Bias, Variance and MSE of the Sampling Distributions of the Estimated $\hat{\beta}_0$, $\hat{\beta}_1$ and p ($\beta_0 = -4.12, \beta_1 = 1.18$)

Design		$\hat{\beta}_0$			$\hat{\beta}_1$			$\hat{p}$	
k	n	Method	Bias	Variance	MSE	Bias	Variance	MSE	MSE
5	10	Berkson	0.5338	0.5516	0.8355	-0.1509	0.0360	0.0587	0.0043
		MLE	0.3071	0.6596	0.7526	-0.0849	0.0425	0.0496	0.0041
	20	Berkson	0.2019	0.5128	0.5526	-0.0589	0.0347	0.0381	0.0030
		MLE	0.0138	0.5185	0.5177	-0.0028	0.0356	0.0355	0.0025
	30	Berkson	0.0761	0.4076	0.4126	-0.0211	0.0261	0.0265	0.0023
		MLE	-0.1781	0.4963	0.5270	0.0516	0.0340	0.0366	0.0019
10	10	Berkson	0.5725	0.3441	0.6712	-0.1635	0.0230	0.0497	0.0031
		MLE	0.2650	0.3776	0.4471	-0.0763	0.0250	0.0308	0.0025
	20	Berkson	0.3041	0.2751	0.3671	-0.0804	0.0195	0.0259	0.0017
		MLE	-0.0125	0.2729	0.2725	0.0016	0.0200	0.0199	0.0011
	30	Berkson	0.1925	0.2288	0.2655	-0.0558	0.0152	0.0183	0.0014
		MLE	-0.0114	0.2157	0.2154	0.0036	0.0156	0.0156	0.0009
15	10	Berkson	0.5722	0.2742	0.6011	-0.1651	0.0167	0.0440	0.0025
		MLE	-0.1637	0.5057	0.5315	0.0438	0.0355	0.0378	0.0020
	20	Berkson	0.3065	0.1949	0.2885	-0.0889	0.0138	0.0216	0.0014
		MLE	0.0197	0.2157	0.2156	-0.0074	0.0154	0.0155	0.0009
	30	Berkson	0.2303	0.1366	0.1894	-0.0652	0.0097	0.0139	0.0009
		MLE	0.0032	0.1221	0.1218	-0.0006	0.0089	0.0089	0.0006

DISTRIBUTION OF THE $\hat{B}_0$ 'S

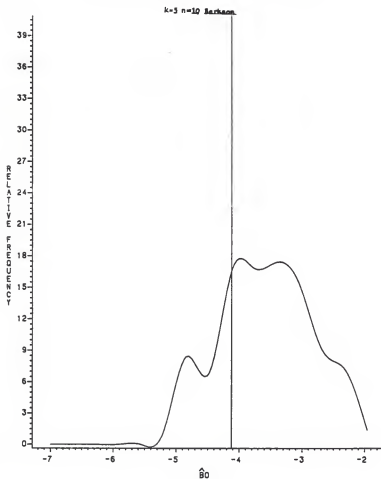


FIGURE 5.1

DISTRIBUTION OF THE $\hat{B}_0$ 'S

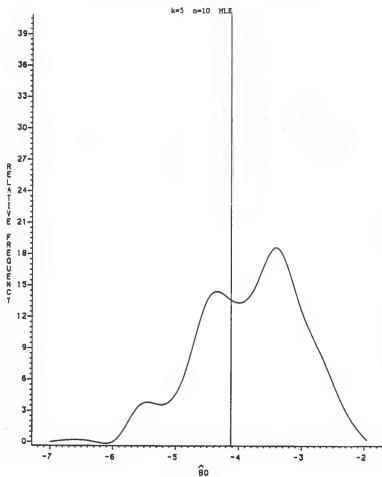


FIGURE 5.2

DISTRIBUTION OF THE $\hat{B}_1$ 'S

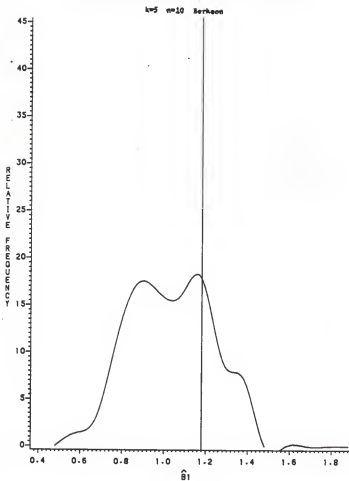


FIGURE 5.3

DISTRIBUTION OF THE $\hat{B}_1$ 'S

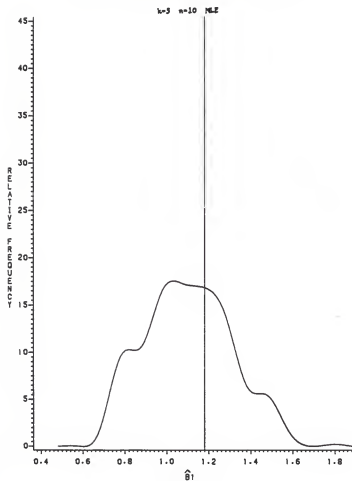


FIGURE 5.4

DISTRIBUTION OF THE $\hat{B}O$ 'S

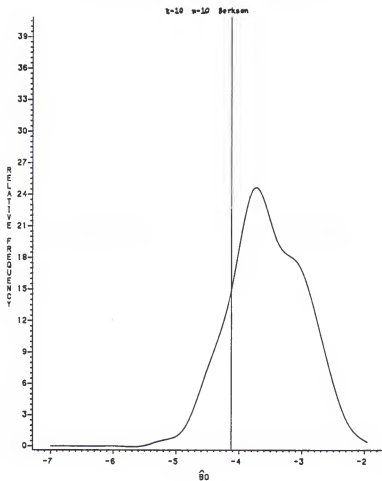


FIGURE 5.5

DISTRIBUTION OF THE $\hat{B}_0$ 'S

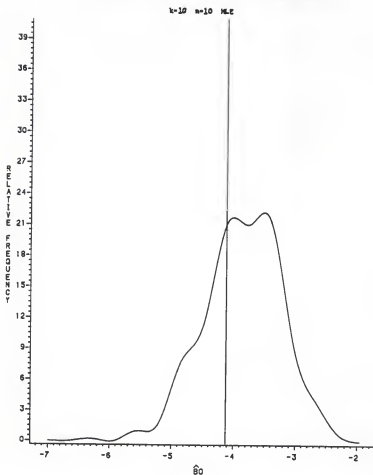


FIGURE 5.6

DISTRIBUTION OF THE $\hat{B}_1$ 'S

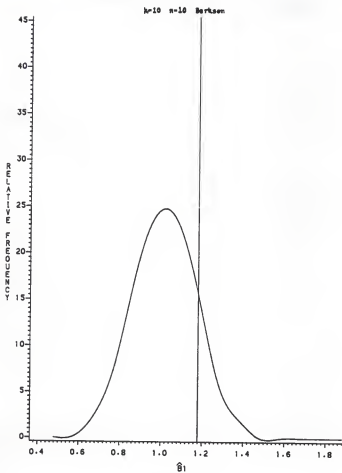


FIGURE 5.7

DISTRIBUTION OF THE $\hat{B}_1$ 'S

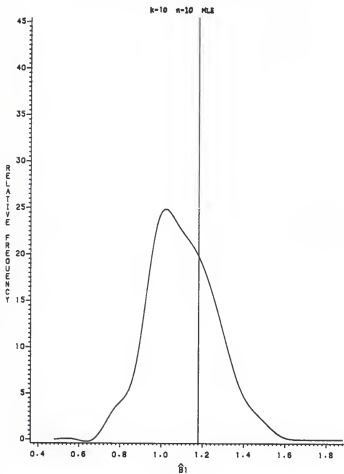


FIGURE 5.8

DISTRIBUTION OF THE $\hat{B}_0$ 'S

k=15 n=10 Berkson

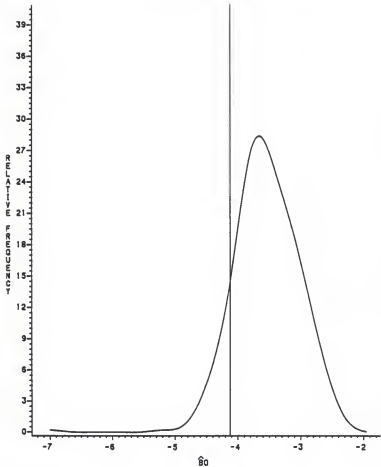


FIGURE 5.9

DISTRIBUTION OF THE $\hat{\theta}_0$ 'S

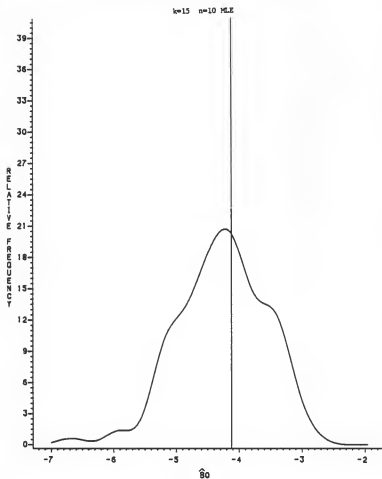


FIGURE 5.10

DISTRIBUTION OF THE $\hat{B}_1$ 'S

k=15 n=10 Berkson

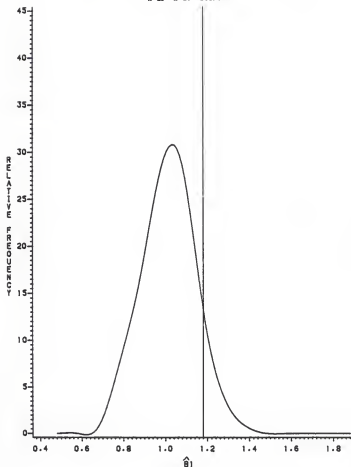


FIGURE 5.11

DISTRIBUTION OF THE $\hat{B}_1$ 'S

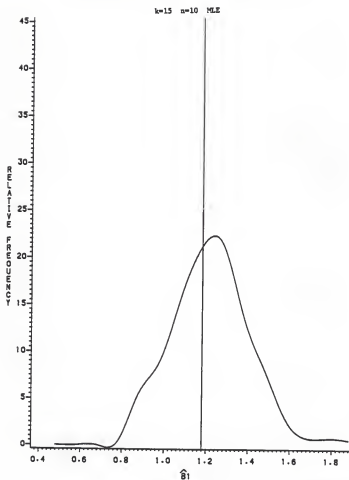


FIGURE 5.12

DISTRIBUTION OF THE $\hat{B}_0$ 'S

k=15 n=30 Berkson

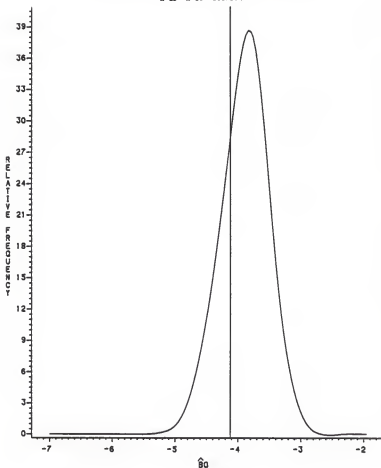


FIGURE 5.13

DISTRIBUTION OF THE $\hat{\beta}_0$ 'S

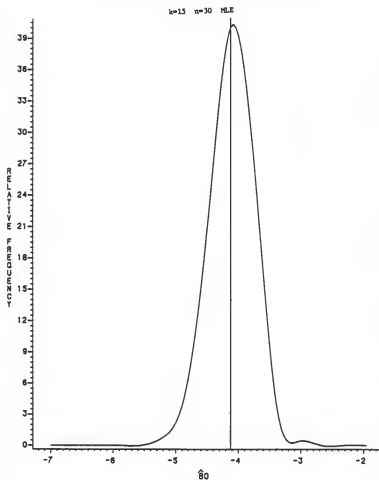


FIGURE 5.14

DISTRIBUTION OF THE $\hat{B}_1$ 'S

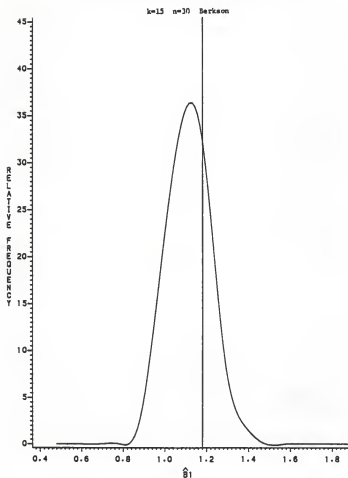


FIGURE 6.15

DISTRIBUTION OF THE $\hat{B}_1$ 'S

$k=15$ $n=30$ MLE

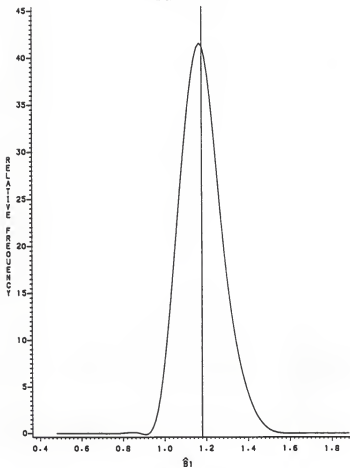


FIGURE 5.16

DISTRIBUTION OF THE $\hat{B}_0$ 'S

k=15 n=20 Berkson

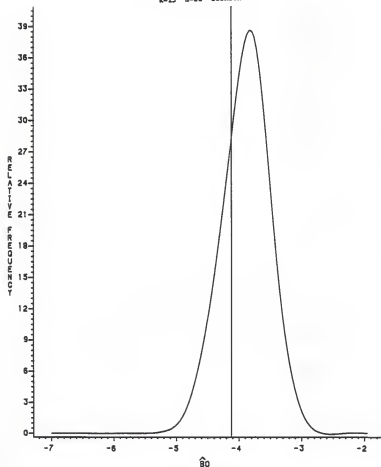


FIGURE 5.17

DISTRIBUTION OF THE $\hat{\beta}_0$ 'S

k=15 n=20 MLE

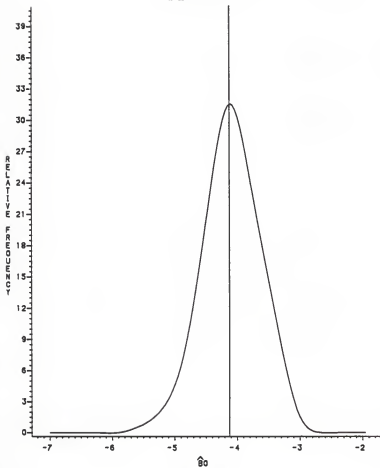


FIGURE 5.16

DISTRIBUTION OF THE $\hat{B}_1$ 'S

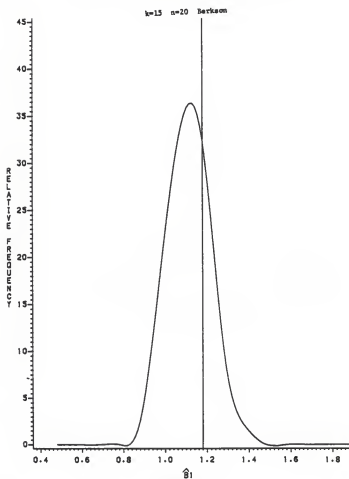


FIGURE 5.19

DISTRIBUTION OF THE $\hat{B}_1$ 'S

k=15 n=20 MLE

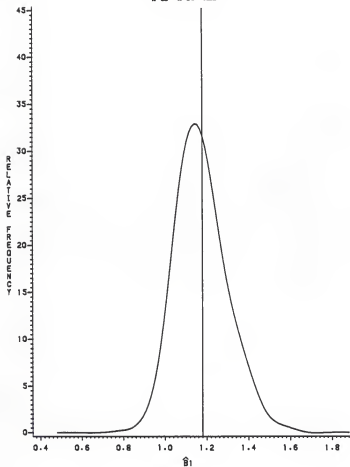


FIGURE 6.20

DISTRIBUTION OF THE $\hat{B}_0$ 'S

k=10 n=30 Berkson

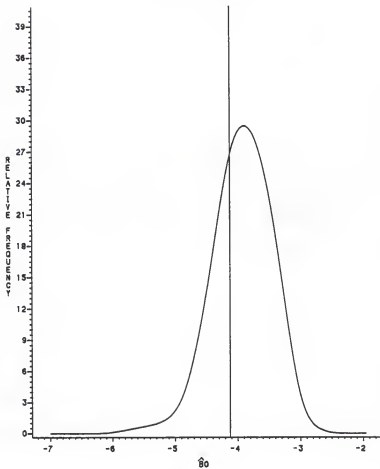


FIGURE 5.21

DISTRIBUTION OF THE $\hat{\beta}_0$ 'S

k=10 n=30 NLE

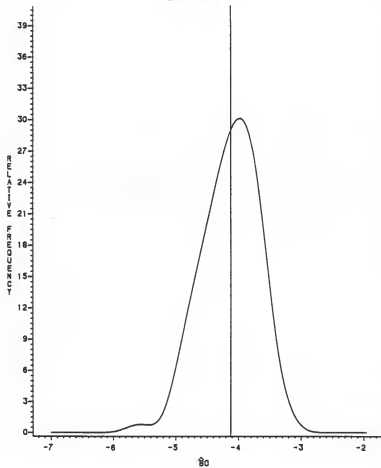


FIGURE 5.22

DISTRIBUTION OF THE $\hat{B}_1$ 'S

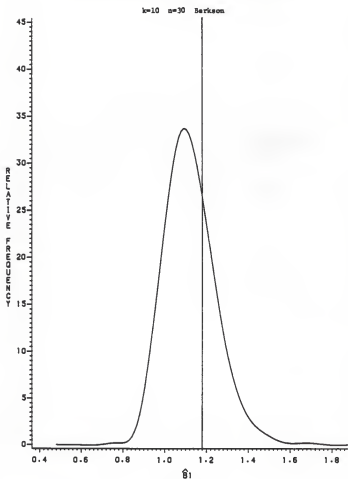


FIGURE 5.23

DISTRIBUTION OF THE $\hat{B}_1$ 'S

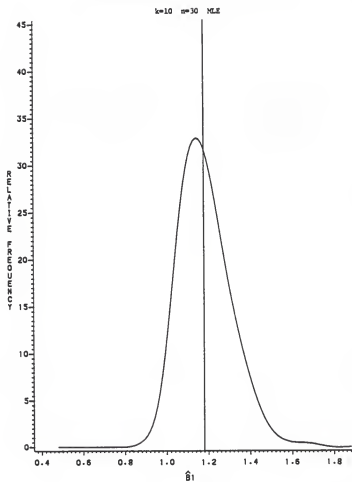


FIGURE 5.24

DISTRIBUTION OF THE $\hat{\theta}_0$ 'S

k=10 n=20 Berkson

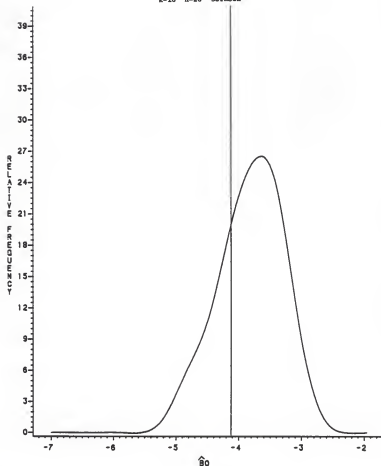


FIGURE 5.25

DISTRIBUTION OF THE $\hat{\beta}_0$ 'S

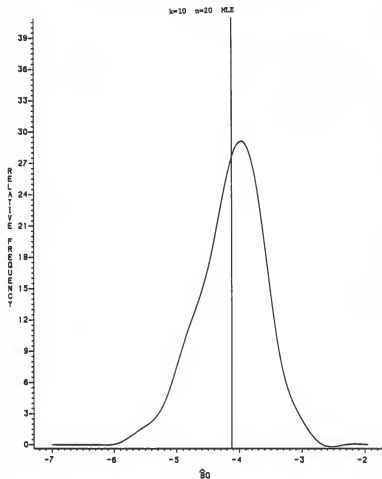


FIGURE 5.26

DISTRIBUTION OF THE $\hat{B}_1$ 'S

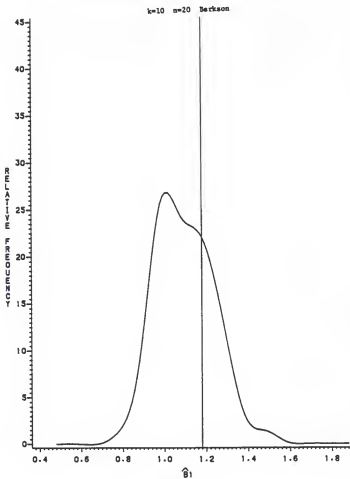
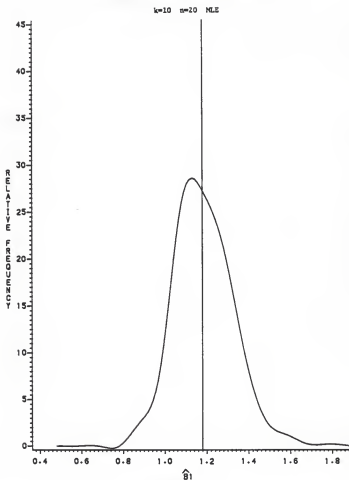


FIGURE 5.27

DISTRIBUTION OF THE $\hat{B}_1$ 'S



DISTRIBUTION OF THE $\hat{B}_0$ 'S

$k=5$ $n=20$ Berkson

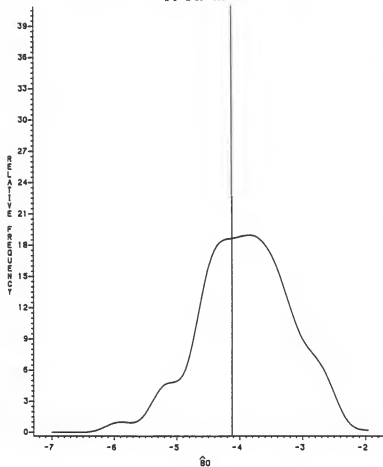


FIGURE 5.29

DISTRIBUTION OF THE $\hat{\theta}_0$ 'S

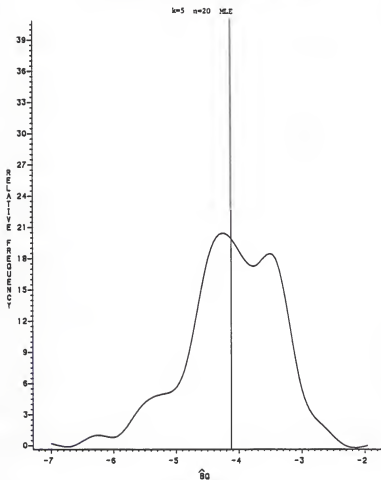


FIGURE 5.30

DISTRIBUTION OF THE $\hat{B}_1$ 'S

$k=5$ $n=20$ Berkson

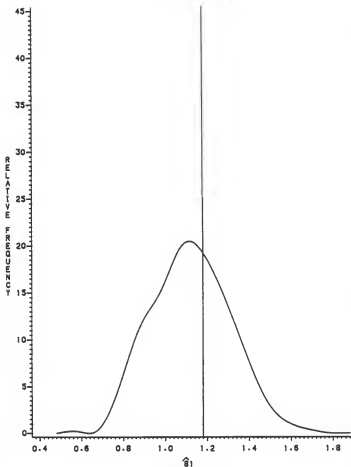


FIGURE 5.31

DISTRIBUTION OF THE $\hat{B}_1$ 'S

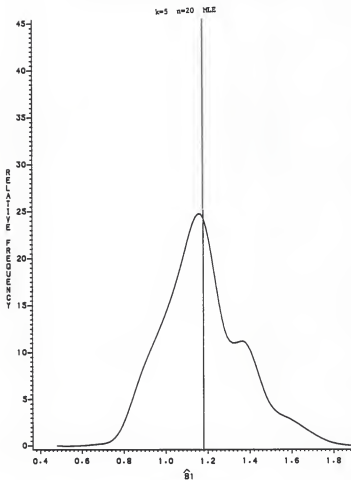


FIGURE 5.3.2

DISTRIBUTION OF THE $\hat{\theta}_0$ 'S

$k=5$ $n=30$ Berkeson

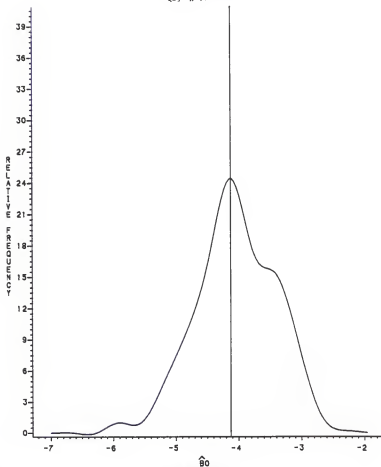


FIGURE 5.33

DISTRIBUTION OF THE $\hat{\theta}_0$ 'S

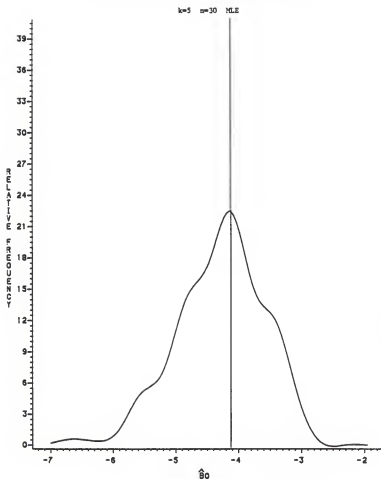


FIGURE 5.34

DISTRIBUTION OF THE $\hat{B}_1$ 'S

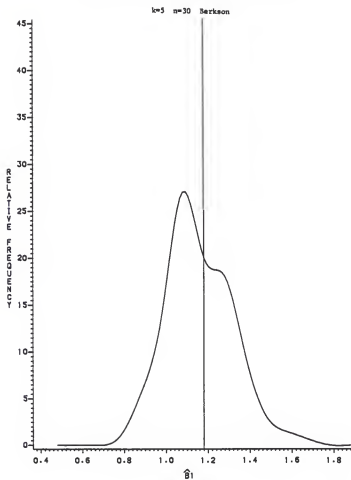


FIGURE 5.35

DISTRIBUTION OF THE $\hat{B}_1$ 'S

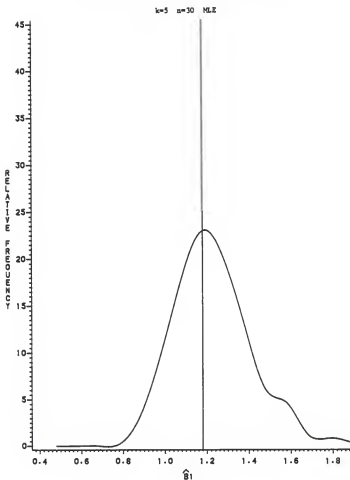


FIGURE 5.36

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=5 n=10 Berkson

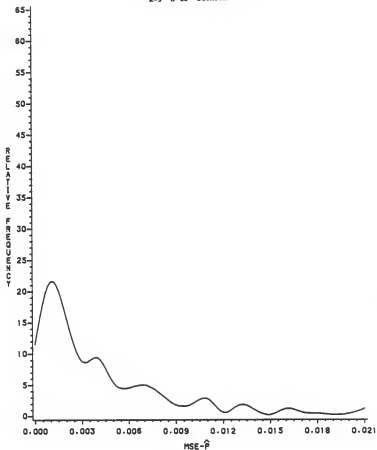


FIGURE 5.37

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=3 n=10 MLE

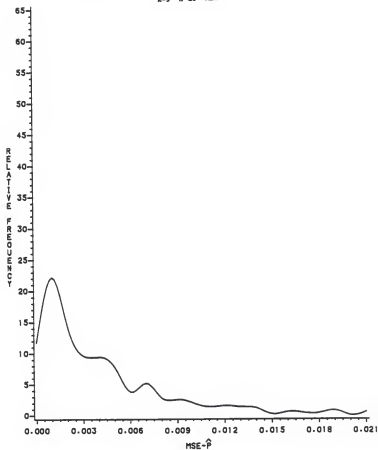


FIGURE 5.38

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=5 n=20 Berkson

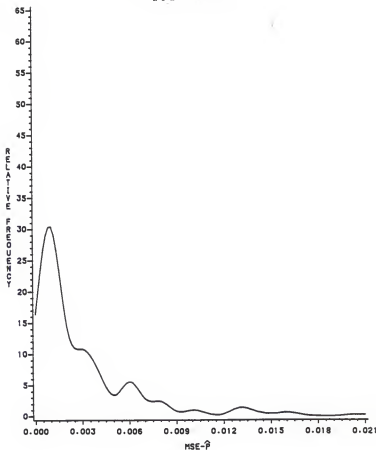


FIGURE 5.39

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=5 n=20 MLE

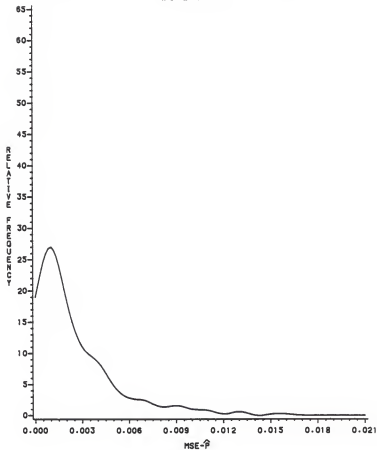


FIGURE 5.40

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=5 n=30 Berkson

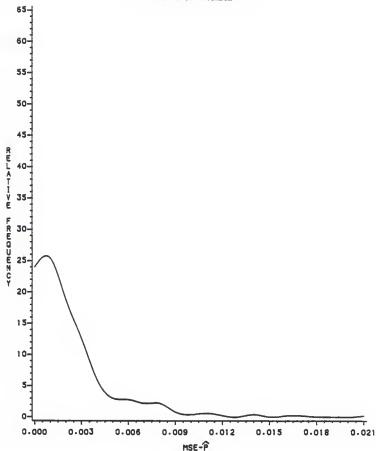


FIGURE 5.41

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=5 n=30 MLE

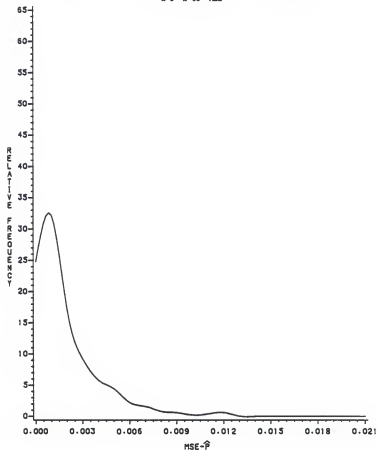


FIGURE 5.42

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=10 n=10 Berkson

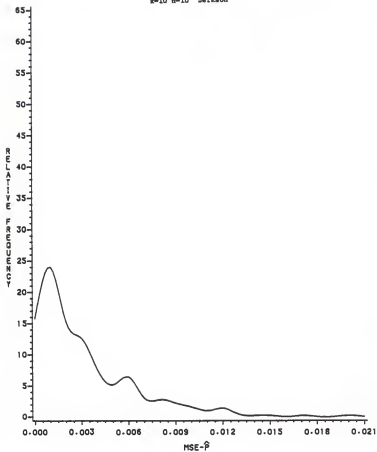


FIGURE 5.43

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=10 n=10 MLE

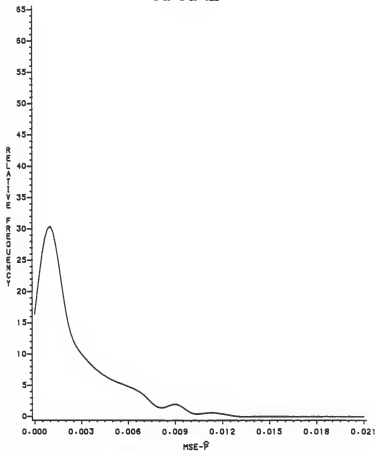


FIGURE 5.44

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=10 n=20 Berkson

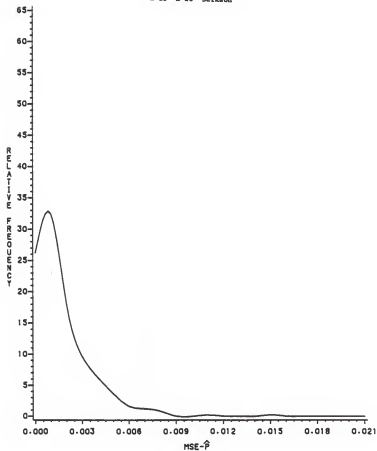


FIGURE 5.45

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=10 n=20 MLE

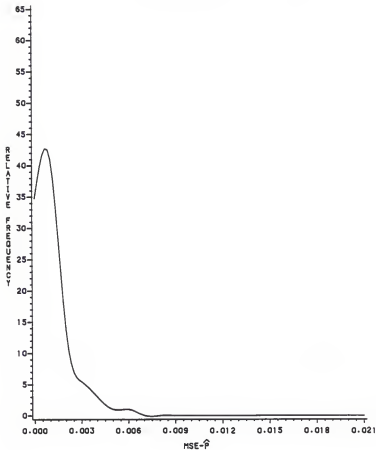


FIGURE 5.46

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=10 n=30 Berkson

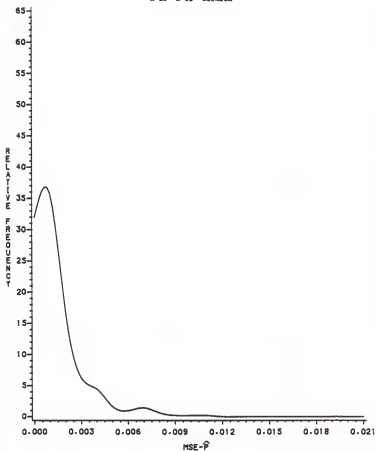


FIGURE 5.47

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=10 n=30 MLE

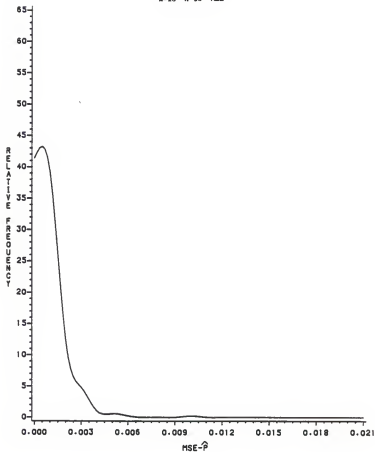


FIGURE 5.48

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=15 n=10 Berkson

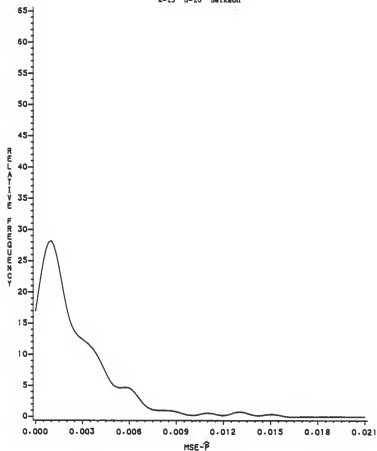


FIGURE 5.49

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=15 n=10 HLE

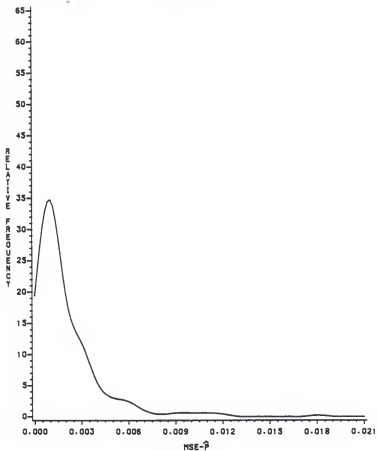


FIGURE 5.60

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=15 n=20 Berkson

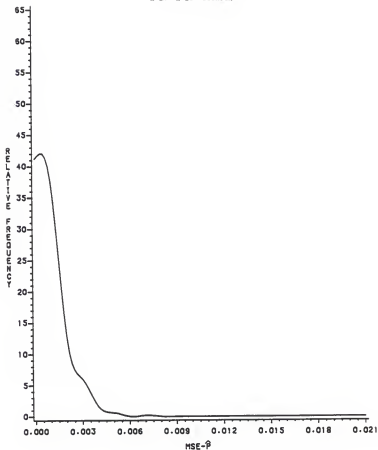


FIGURE 5.51

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=15 n=20 MLE

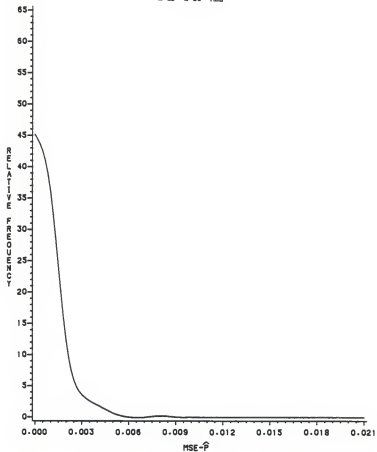


FIGURE 5.52

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

k=15 n=30 Berkson

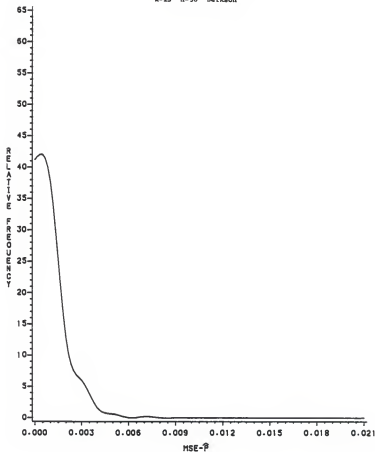


FIGURE 5.53

DISTRIBUTION OF THE MSE $\hat{P}$ 'S

$k=15$ $n=30$ MLE

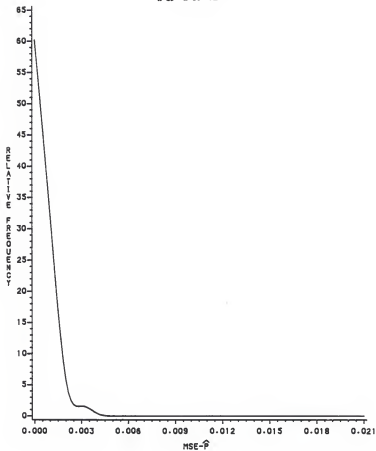


FIGURE 5.54

VI. CONCLUSIONS AND RECOMMENDATIONS

These simulation results concur with the conclusion by Davis(1985) that Berkson's method is not advisable when the number of design points(k) is large but the average number of observations(n) per design point is small.

The maximum likelihood demonstrated superiority in every criterion except for the variances under all designs but one(5 dose, 30 observations per dose case). Most importantly, it fared consistently better in the $MSE(\hat{p})$ criterion which is probably of utmost consideration in practical applications.

Although the maximum likelihood is the method recommended for use based on the results of this study, the simulation results did show some indication that Berkson's method may be better for experiments when the number of design points is small but the number of observations per design point is large. Further studies may be in order regarding that situation.

VII. BIBLIOGRAPHY

- Berkson, Joseph, "Application of the Logistic Function to Bio-Assay",
Journal of American Statistical Association, 39(1944), 357-365.
- Berkson, Joseph, "A Statistically Precise and Relatively Simple
Method of Estimating the Bioassay with Quantal Response, based on
the Logistic Function", Journal of American Statistical
Association, 48(1953), 565-599.
- Berkson, Joseph, "Maximum Likelihood and Minimum X^2 Estimates",
Journal of American Statistical Association, 50(1955), 130-162.
- Chatterjee, S. and Bertram Price, Regression Analysis by Example,
Wiley, New York, 1977.
- Davis, Linda J., "Consistency and Asymptotic Normality of the Minimum
Logit Chi-Squared Estimator when the number of Design Points is
Large", The Annals of Statistics, 13(1985), 947-957.
- Edlefsen, L. and Samuel D. Jones, GAUSS Programming Language Manual,
WA:Aptech Systems, Inc., 1986.
- Finney, D.J., Statistical Method in Biological Assay, Hafner, 1971.
- Milliken, George A., Nonlinear Statistical Models, Unpublished, 1982.
- SAS Institute Inc., SAS User's Guide:Statistics, 1982 Edition, Cary,
NC: SAS Institute Inc., 1982.
- Serfling, Robert., Approximation Theorems of Mathematical Statistics,
Wiley, 1980.

VIII. APPENDIX

A. Sample Gauss Program : Generation Of Data Sets

```

seed=45989;
i=451;
do while i <=500;
  dose=1;
  do while dose <= 6;
    pow=-4.12 + (1.18*dose);
    pt=1/(1+exp(-pow));
    rannum=rndus(20,1,seed);
    dead=rannum[.,1]< pt;
    dnum=sumc(dead);if dnum=0;dnum=20*(1/40);
    endif;
    if dnum=20;dnum=20*(1-(1/40));
    endif;
    if dose = 1 ; deaths=dnum;x=dose;totl=20;
      ptrue=pt ; sampleno=i;
    else; deaths =
      deaths+dnum;x=x+dose;totl=totl+20;
      ptrue=ptrue|pt;sampleno=sampleno|i;
    endif;
    dose=dose+.555;
  endo ;
enddo ;

/* to form the dose response matrix x */

dataset = x-totl-deaths-ptrue-sampleno;
format /ml/rd 11,8;
output file=dat1020j on; dataset;
i=i+1;seed=seed+100;
enddo;

```

B. Sample Gauss Program : Berkson Estimates

```

/* Berkson : this program computes Berkson estimates */

p=deaths./totl;
pprime=ln(p./(1-p));
weight=totl.*p.*(1-p);
wpprimex=weight.*pprime.*x;
wprim=weight.*pprime;
wx=weight.*x;
wxsqr=weight.*(x.*x);
bl=((sumc(wpprimex))-(((sumc(wprim))*sumc(wx))/sumc(weight))))/
    ((sumc(wxsqr))-(((sumc(wx))^2)/sumc(weight))));
bo=(meanc(pprime))-(bl*meanc(x));

/* to compute the mse */

phat=(exp(bo+(bl*x)))/(1+(exp(bo+(bl*x))));
ssep=meanc((phat-pttrue)^2);
ssebo=(bo+4.12)^2;
ssebl=(bl-1.18)^2;

/* to output into a data set */

    if i==1 ; matbl=bl;matbo=bo;matp=ssep;matsbl=ssebl;matsbo=ssebo;
        else;
matbl=matbl|bl;matbo=matbo|bo;matp=matp|ssep;matsbl=matsbl|ssebl;
        matsbo=matsbo|ssebo;
        endif;
seed=seed+100;
i=i+1;

endo;

/* to print the matrix */

est=matbl-matbo-matp-matsbl-matsbo;
format %ml/%rd 11,8;
exvalue=meanc(est);dev=stdc(est);summary=exvalue-dev;
lprint "ESTIMATES AND MEAN SQUARE ERRORS(CHIT1020)" est;
output file=b:chit1020 on;
lprint "SUMMARY STATS(chit1020)" summary';

```

C. SAS Program : Maximum Likelihood Estimates

```

*-----*
*          THIS PROGRAM COMPUTES THE          *
*          MAXIMUM LIKELIHOOD ESTIMATES       *
*          FOR THE LOGIT MODEL                *
*-----*
%MACRO ANALYSIS;
  %DO I=1 %TO 500;
    DATA ANA&I; SET RJR.SIM1520;
    IF SAMPLENO=&I;
    DATA ANA2&I; SET ANA&I;
    COUNT=DEATHS; Y=1; OUTPUT;
    COUNT=TOTAL-DEATHS; Y=0; OUTPUT;
    PROC NLIN NOHALVE SIGSQ=1 /*ALGORITHM FROM SAS82 STATISTICS
P36*/;
      PARMS BO=-4.6 B1=1.3 /*INITIAL ESTIMATES OF BO AND B1 FROM
PROC REG*/;
      E=EXP(BO + B1*DOSE);
      P=E/(1+E);
      MODEL Y=P;
      W=1/(P*(1-P));
      WEIGHT =W*COUNT;
      DER=E/((1+E)**2);
      DER.BO=DER;
      DER.B1=DER*DOSE;
      OUTPUT OUT=RESULTS P=PHAT PARMS=BO B1;
    DATA PREFINAL;
      KEEP BO B1 SSE;
      SET RESULTS;
      SSE=((PHAT-PTURE)**2);
    PROC MEANS MEAN NOPRINT;
      OUTPUT OUT=FINAL MEAN=BO B1 MSE;
    PROC APPEND BASE=FINALDAT DATA=FINAL;
  %END;
%MEND;
%ANALYSIS;
DATA RJR1.ML1520; SET FINALDAT;

```

BERKSON'S METHOD VS. MAXIMUM LIKELIHOOD
IN THE LOGISTIC REGRESSION MODEL

by

RHONDA JENINE UMALI RAGUAL
B.S., UNIVERSITY OF THE PHILIPPINES AT LOS BANOS, 1983

AN ABSTRACT OF A MASTERS REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY
Manhattan, Kansas
1988

ABSTRACT

Two estimators(Maximum Likelihood and Berkson's(1953)) of the logistic regression model have been proposed in the literature. The two estimators were compared by simulation under 9 different dose-response designs. Maximum likelihood emerged as the better estimator in all cases considered except one. Berkson's method showed some promise in a few design point-many observations per design point set-up.